

Increasing reality of species distribution models of consumers by including its food resources

Gabriel Preuss¹, André Andrian Padial¹

¹ *Laboratório de Análise e Síntese em Biodiversidade, Departamento de Botânica. Programa de Pós-graduação em Ecologia e Conservação, Universidade Federal do Paraná, 81531-980, Curitiba-PR, Brazil*

Corresponding author: André Andrian Padial (aapadial@gmail.com)

Academic editor: A.M.Leal-Zanchet | Received 23 February 2021 | Accepted 7 July 2021 | Published 6 August 2021

Citation: Preuss G, Padial AA (2021) Increasing reality of species distribution models of consumers by including its food resources. *Neotropical Biology and Conservation* 16(3): 411–425. <https://doi.org/10.3897/neotropical.16.e64892>

Abstract

Species distribution models are not usually calibrated with biotic predictors. Our study question is: does the use of biotic predictors matter in predicting species distribution? We aim to assess the importance of biotic predictors in the output of distribution models of the Brazilian squirrel (*Sciurus aestuans*) throughout South America based on fruits of *Syagrus romanzoffiana* – the most consumed food resource. We hypothesized that the distribution model of *S. aestuans* using its main food resource as a biotic predictor will be more accurate in comparison with the output of the model without the biotic predictor. We built three different distribution models: (i) distribution of *S. romanzoffiana*; (ii) distribution of *S. aestuans* without biotic predictor; and (iii) distribution of *S. aestuans* with biotic predictor. We evaluated performance scores, number of presence pixels and concordance between suitability maps. We found that performance scores may not vary between models with different predictors, but the output map changed significantly. We also found that models with biotic predictors seem to vary less in presence pixels. Furthermore, the main variable in the distribution model was the biotic variable. We conclude that the knowledge of a species' biology and ecology can make better predictions of species distribution models mainly by avoiding commission errors.

Keywords

biotic interaction, biotic predictor, diet specialization, Eltonian Noise Hypothesis, hutchinsonian niche, realized niche, variables importance

Introduction

One of the most important characteristics of species distribution modelling (SDM) is the capacity to generate accurate predictions of past, present and future species distributions (Guisan and Thuiller 2005; Barry and Elith 2006). Assuming that models have a minimally reliable predictive capacity, conservation planning can be elaborated with support of these findings, especially in a changing world (Margules and Usher 1981; Haila and Kouki 1994). Historically, SDMs are calibrated only with abiotic variables as predictor, assuming that biotic interactions do not influence species range pattern, have influence only at small spatial scales, or are indirectly represented by abiotic variables (Huntley et al. 1995; Bakkenes et al. 2002; Pearson and Dawson 2003; Heikkinen et al. 2007). In fact, the Eltonian Noise Hypothesis (ENH) assumes that biotic interactions do not influence species distribution (Soberón and Nakamura 2009). Modeling exercises using abiotic features thus consider species distribution inside the Grinnellian niche, which comprises the scenopoetic variables (Grinnell 1917). Even so, some studies suggest the important role of biotic interactions as a predictor in spatial distribution even on a large scale (Meier et al. 2011; Araújo et al. 2014). By including variables representing biotic interactions, modeling exercises can thus assume a Hutchinsonian view of the niche: predation, competition, facilitation and mutualism together with environmental filtering compose the realized niche (Hutchinson 1957).

One difficulty to insert biotic interactions in SDMs is the lack of suitable data to represent biotic interaction on a large-scale (Araújo and Luoto 2007; Zimmermann et al. 2010; Guillera-Aroita et al. 2015), especially when species interactions are poorly known or its interactors are numerous (Meier et al. 2011). Since occurrence data are increasing in open access databases, obtaining large-scale data has become easier and cheaper (Fielding and Bell 1997; Wisz et al. 2013). In this way, studies that use a well-distributed species with sufficient sampling effort and its biotic interactions may fill the information gap in biotic interactions' role in SDMs (Meier and Dikow 2004).

Indeed, robust models require two principal components: sufficient sampling effort in occurrence data for the target biological group and a set of meaningful predictor variables with low or no collinearity (Lobo et al. 2002; Synes and Osborne 2011). There are different ways to assess the performance of models, and the most common are model features such as prevalence, Kappa, True Skill Statistic and AUC scores (Allouche et al. 2006; Liu et al. 2011). Performance scores aims to measure the accuracy of species distribution models through discrimination capacity and reliability, and there are different applications for each type of performance score (Pearce and Ferrier 2000; Liu et al. 2011). Those scores can be divided in three: (i) Threshold-independent scores are those which can be assessed with continuous values (suitability models); (ii) Threshold-dependent are scores only assessed with binary data (presence-absence models); and (iii) overall ones are those applied to both cases (Liu et al. 2011).

The assessment of a model's performance may provide the possibility to compare models with alternative algorithms and predictors and evaluate how different variables affect the model's predictive performance (Kadmon et al. 2003; Loiselle et al. 2003; Segurado and Araujo 2004; Pearson et al. 2007). Furthermore, each study area may have its particular source of variation. In North America, for example, the variation in final results of SDM exclusively by algorithm may be high, while in South America the variation reaches lower oscillations (Diniz-Filho et al. 2009), suggesting that for South America, the studies might focus on better predictor variables instead of considering algorithm variation.

Keeping the focus on biotic predictors, a previous study that incorporated competition in predictor variables showed that biotic interactions influence directly the predicted range of tree species (Meier et al. 2011). These findings suggest that one species may be determinant in another species distribution. It is known that the addition of biotic interactions usually improves the predictive performance of SDMs (Araújo and Luoto 2007). Furthermore, the inclusion of biotic interactions in SDMs has restricted the predicted distribution of a species in relation to another (Schweiger et al. 2012), reducing the possibility of commission error (false positives). However, the inclusion of biotic interactions in SDMs has been concentrated in competition (Guisan and Thuiller 2005; Meier et al. 2011). Other kinds of biotic interactions such as predation, herbivory and facilitation still need to be tested (Guisan and Thuiller 2005; Baselga and Araújo 2009).

One way to guarantee that one kind of interaction really occurs in nature, is the dependency of the species with that interaction (Meier et al. 2011), a feature that usually occurs with food specialist species. In this way, the relation resource-consumer may be a good way to understand the role of biotic interactions in species distributions; furthermore, this relation was recorded as necessary additional information for the knowledge of the role of biotic interactions in SDMs (Anderson 2017). Herein we used a food specialist species of squirrel to predict its actual distribution once its distribution is poorly known in Brazil, but its foraging behavior is marked by the strong interaction with a palm tree fruit (Alvarenga and Talamoni 2006).

Sciurus aestuans Linnaeus, 1776 is a squirrel species widely distributed along South America (Amori et al. 2013). In spite of the generalism in squirrels' diet, in the Brazilian Atlantic Forest, the foraging behavior of *S. aestuans* is characterized by the use of the fruit of *Syagrus romanzoffiana* (Cham.) Glassman (Paschoal and Galetti 1995; Bordignon and Monteiro-Filho 2000; Alvarenga and Talamoni 2006). In a study developed in the Brazilian Atlantic Forest, *S. romanzoffiana* presented up to 70% of *S. aestuans* diet (Alvarenga and Talamoni 2006). The same study shows that the palm tree is essential for the reproduction of squirrels, where lactating females spend their time feeding and not searching for food, suggesting elevated habitat fidelity near the palm tree (Alvarenga and Talamoni 2006). Furthermore, the caching habit of squirrels allows them to consume *S. romanzoffiana* fruits throughout the year, showing a strong interaction between the species (personal observations; Paschoal and Galetti 1995; Bordignon and Monteiro-Filho 2000; Alvarenga and

Talamoni 2006). Considering all this, we believe *S. aestuans* and *S. romanzoffiana* to have a close relationship in relation to ecology and spatial distribution.

Assuming that the distribution pattern of *S. aestuans* is regulated by a bottom-up control, we hypothesized that: I – The distribution model of *S. aestuans* with *S. romanzoffiana* as biotic predictor will be more accurate in comparison with the output of the model without biotic predictor, and II – The Eltonian Noise Hypothesis is not determinant in our distribution models once biotic interaction will affect the distribution area.

Methods

Data preparation

A recent taxonomic review suggests the division in *Sciurus* genus. However, in GBIF database the occurrence points are available for *S. aestuans*. To obtain the occurrence records of *S. aestuans* and *S. romanzoffiana*, we used personal data and the Global Biodiversity Information Facility (GBIF) database (GBIF 2020), with data recorded from years 1979 to 2013. We selected the data for South America, which resulted in 149 occurrence points of *S. romanzoffiana* and 132 of *S. aestuans* before data cleaning (presence records of *S. romanzoffiana* and *S. aestuans* can be seen in Suppl. materials 1, 2, respectively).

The environmental variables of each model were obtained from Climatologies at High Resolution for the Earth's Land Surface Areas (CHELSA) database (Karger et al. 2017), collecting all the 19 bioclimatic variables under 30 sec arc for the same period of occurrence records (1979–2013). Since climatic data is available for the whole world, but our analysis was in South America, we cropped the raster files to our study area. Hence, for data cleaning we excluded points inside the same pixel, and those out of the confirmed distribution of the species following a Brazilian checklist established by specialists by *S. romanzoffiana* (Soares, 2020) and out of the boundaries of South America for *S. aestuans* where there is no information about its distribution. We were left with 136 occurrence points of *S. romanzoffiana* and 74 of *S. aestuans*. To avoid collinearity among variables, which would overfit the model, we used Spearman correlation and excluded the variables most correlated by a threshold of 0.6 (Loiselle et al. 2003). Only six variables were retained (mean diurnal range, isothermality, max temperature of warmest month, mean temperature of wettest quarter, precipitation seasonality, precipitation of warmest quarter and precipitation of coldest quarter). To finish data preparation, we ran a background selection with the same number of occurrence points in the entire area of distribution of the species.

Models

Our first model aims to predict the distribution of *S. romanzoffiana* that will be used as predictor variable of *S. aestuans*. Our second model predicts the distribution of *S. aestuans* only considering climatic predictors. Finally, our third model comprises

the predicted distribution of *S. aestuans* with climatic and biotic predictor. The last model is built with the predictors of model 2 (mean diurnal range, isothermality, max temperature of the warmest month, mean temperature of wettest quarter, precipitation seasonality, precipitation of warmest quarter and precipitation of coldest quarter) plus the distribution of *S. romanzoffiana* as a biotic predictor.

Modelling and validation

Considering that the data selected for this study was presence data, we used the Maximum Entropy Model available in *dismo* package in R (Hijmans et al. 2011). Furthermore, this is a model of presence-background, which according to Elith and Graham (2009) performs better than some presence-absence models. Moreover, we emphasize that our purpose is not to find the best algorithm model for the above-mentioned species, but to compare performance of the model with and without the variable representing resources for the consumer. For each model we used 75% of location points as training data and 25% for test validation of the Maxent model. Also, for each model we saved suitability and presence/absence (using the threshold which gets higher specificity and sensibility) raster files, values of variables contribution, values of AUC, kappa and TSS which was built according to the best threshold following the ROC-AUC curve.

With those saved files, we made ensembles with 10 simulations for each model and built the final raster of suitability and presence/absence. We also built a table with the values of model performance (AUC, kappa, prevalence, TSS) for each simulation from models 2 and 3 in order to compare the model performance with and without biotic predictor. Afterward, we made the multiplication of ensembles to assess the concordance of species suitability between maps with and without biotic predictor. Using this method, we can assess the area with greater probability for *S. aestuans* to occur. All codes used may be seen in Suppl. material 3.

Results

Our maps show that the distribution of *S. aestuans* differs according to predictor variables used (Figure 1). For scores of model performance, there was no difference between models without and with biotic predictor use considering standard deviation (Table 1).

Apart from the metrics of performance, models with biotic predictor show an area of suitability smaller than that one without biotic predictor (Fig. 1A, B respectively). Furthermore, looking for variables' contribution of the model, the presence of *S. romanzoffiana* (the resource used as biotic predictor) was the most important variable in species distribution (Fig. 2). Variables contribution and performance scores of all simulations can be seen in Suppl. material 4: Tables S1, S2.

To represent the difference between maps without and with biotic predictors in the same figure, we multiplied the pixel values of both maps (Fig. 3). There was

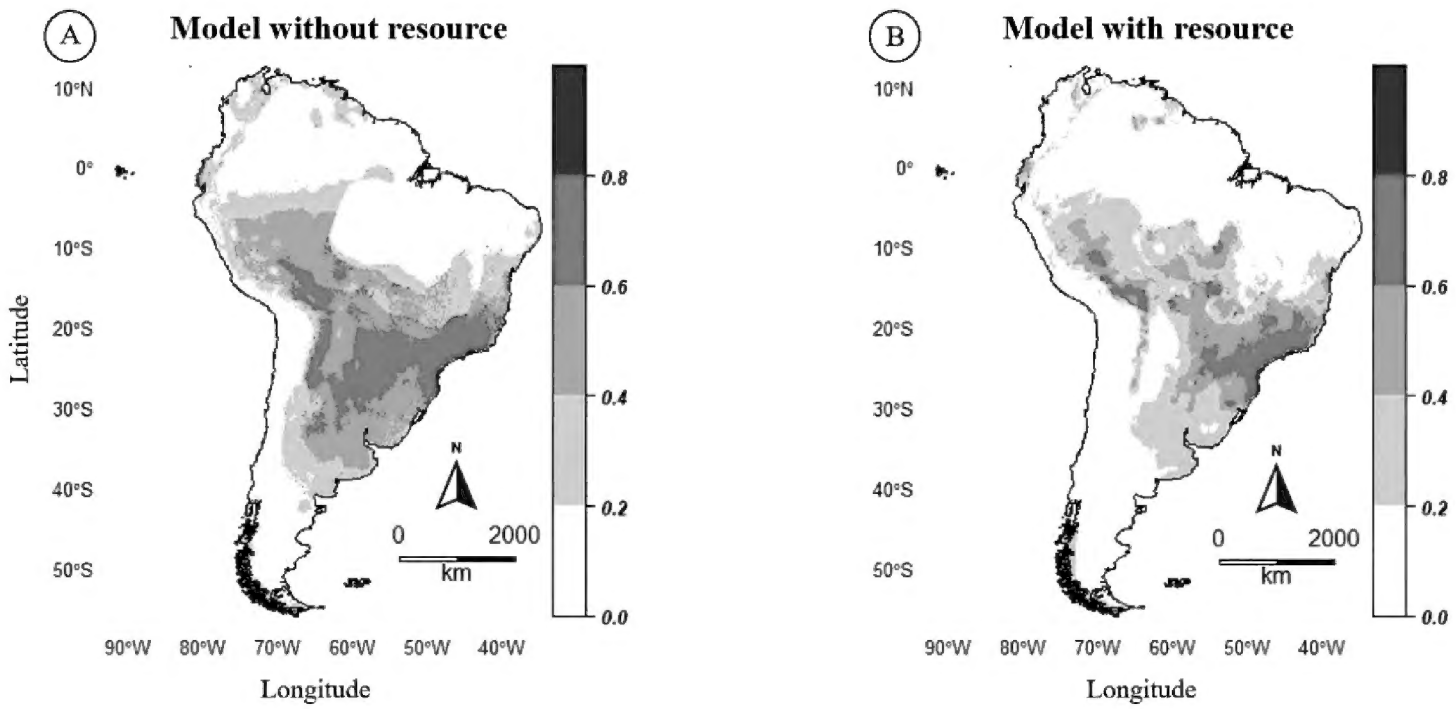


Figure 1. Ensemble from cross validation maxent maps of suitability, higher values present higher habitat suitability and higher chance of species occurrence, where: (A) model without resource presents only climatic predictor variables, and (B) model with resource presents climatic variables plus distribution of *Syagrus romanzoffiana*.

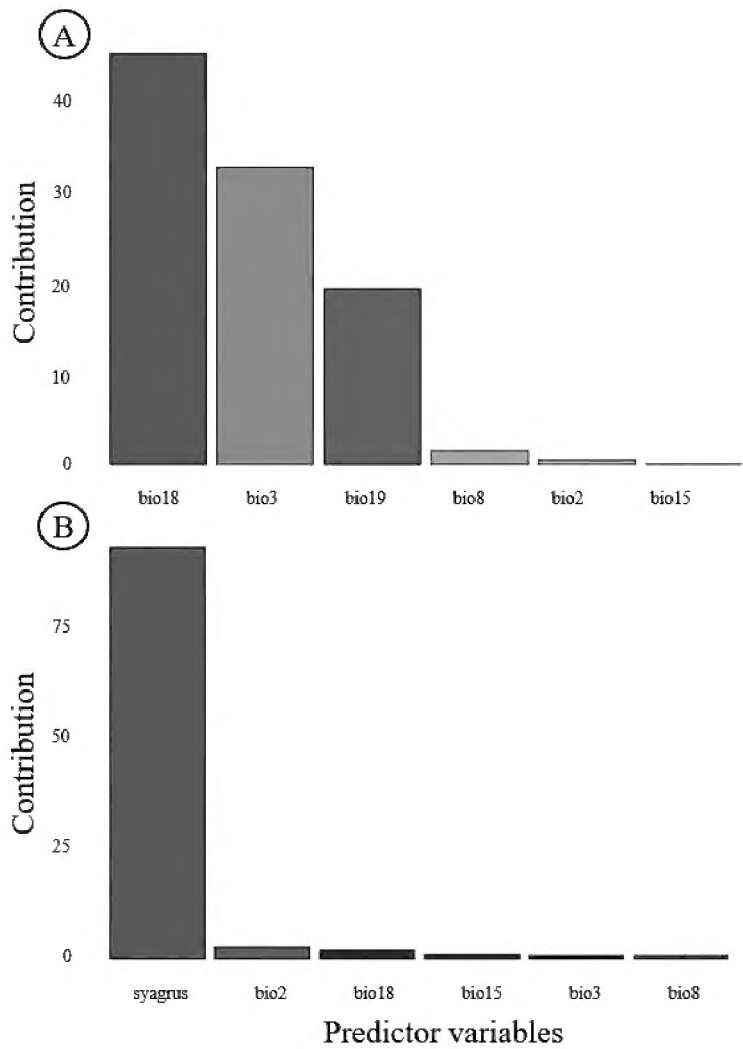


Figure 2. Variable contribution of models without and with biotic interaction, where: (A): contribution of variables in species distribution suitability without biotic predictor and (B): contribution of variables in species distribution suitability with biotic predictor. Predictor variables represents: bio2 = Mean diurnal range, bio3 = Isothermality, bio8 = Mean temperature of wettest quarter, bio15 = Precipitation seasonality, bio18 = Precipitation of warmest quarter, bio19 = Precipitation of coldest quarter, syagrus = predicted distribution of *Syagrus romanzoffiana*.

Table 1. Results of each model performance scores inside cross-validation (n=10) and number of presence pixels for models without and with biotic predictors. Values next to 1 present higher model accuracy and SD represents Standard Deviation.

| Values | Kappa | AUC | TSS | Presence pixels |
|------------------------------------|-------|-------|-------|-----------------|
| Model with biotic predictor | | | | |
| Mean | 0.527 | 0.983 | 0.527 | 2819726 |
| SD | 0.069 | 0.013 | 0.069 | 1518457 |
| Without biotic predictor | | | | |
| Mean | 0.531 | 0.991 | 0.531 | 4660709 |
| SD | 0.053 | 0.007 | 0.053 | 1939159 |

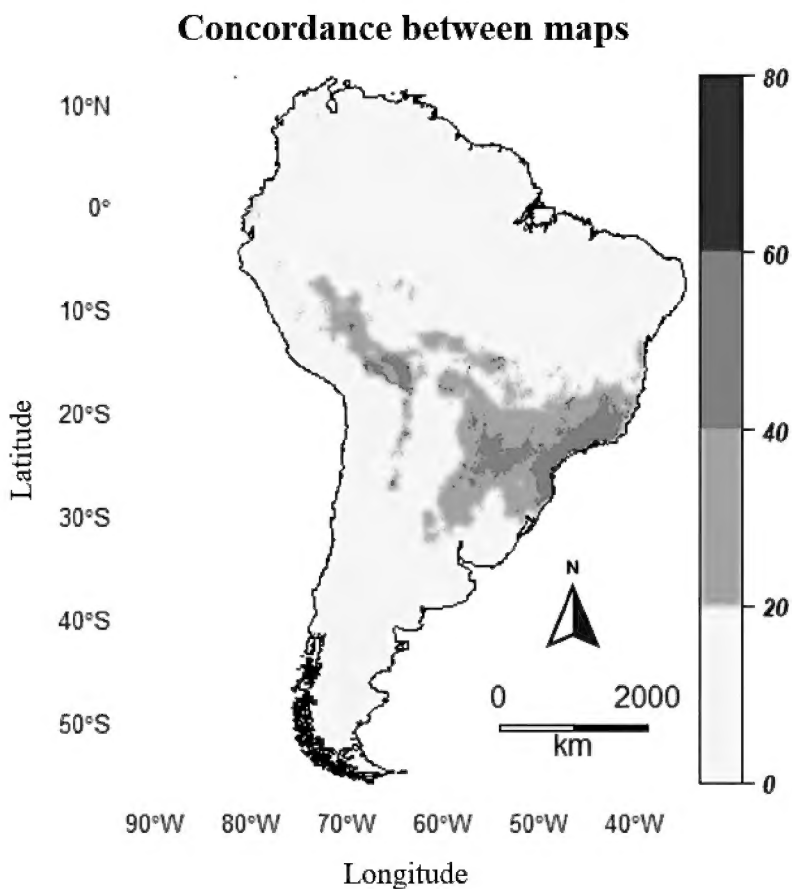


Figure 3. Output of concordance between maps with and without biotic predictor. Values above 20 and under 40 indicate high suitability for the species only in the model without biotic predictor. Values above 40 indicate high suitability in both models.

concordance between models without and with biotic predictors. However, a model without biotic predictor suggests a larger possible occurrence area for the species. Besides, the variation in presence pixels among the models of ensemble is higher without biotic predictor (Suppl. material 4: Figure S1).

Discussion

The results indicate that our hypotheses are not rejected. The importance of including a variable representing a biotic interaction was clear in our modeling exercise. Suitability maps were more restricted by the biotic predictor, which was the most determinant variable in the distribution model. This indicates that commission errors were probably reduced considering the predominant food resource of

the consumer. Concerning performance scores, both models seem to be similar. This means that model refinement considering biotic interactions did not affect the performance of the modeling exercise. However, practical results such as the presence/suitability area for species distribution suggest the contrary. This leads to an important finding: models can reach high values of accuracy in scores but represent distinct results in terms of species distribution. In this way, we suggest that before making a modeling exercise, one should ask: do most used performance scores truly represent the reality and reliability of a model?

Even though the insertion of a biotic predictor may lead to neutral or worse model performance (Araújo and Luoto 2007), our models performed similarly. We used the three types of performance scores: threshold-independent, threshold-dependent and overall test of SDM performance (AUC, Kappa and TSS respectively) to compare the difference in score values (Liu et al. 2011). However, there was no difference in performance score values, which may tell us that they should not be considered alone. More important than selecting the best variables for prediction, we suggest that researchers should question themselves if the most important variables really matter considering the biology of the target species.

Furthermore, comparing the variation in presence pixels of each simulation inside the ensemble, the model with the biotic predictor tends to be more accurate and closer to the final ensemble. The lower variation may also indicate lower commission errors (Peterson et al. 2008). In addition, the resulting suitability map changed significantly, where the model with the biotic predictor presents an area of presence pixels that is half the size of the model without the biotic predictor. Considering that for conservation priority areas we cannot reach all areas and need to choose those which concentrate the higher probability of species occurrence, practical results may be more, or as crucial as, model performance. In this way, what can make the difference to prioritize an area or not are the importance of predictors used. Therefore, biotic predictors may not significantly influence performance scores, but certainly influence the result of species distribution area.

The fact that the biotic predictor was by far the most important indicates that biotic interactions may have a central role in species distribution, differently from what is expected in ENH. Surely, we cannot discard the possibility that the correlation of occurrences in data from the squirrel and from the fruit can be due to a sampling bias – researchers sample the same grids. We recommend that this should be further investigated. Even so, inventories are not usually done simultaneously by researchers from different groups and we then raised the possibility that results highlight the importance of knowing the biology of the modelled species. Notwithstanding the controversy of considering biotic predictors as determinant in species distribution, this seems to be a pattern corroborated by several studies (Meier et al. 2011; Meineri et al. 2012; Araújo and Rozenfeld 2013; Giannini et al. 2013; Palacio and Girini 2018; Simoes and Peterson 2018). Such a pattern seems to repeat itself especially for endotherms, since the distribution of ectotherms may be directly related

to climate conditions instead of biotic interactions, and the addition of biotic predictors may increase model complexity and decrease model performance (Engler et al. 2013), unless species have a strong biotic interaction. Anyway, we advocate that the knowledge of species interaction is essential for SDMs.

Since SDMs have been the most applied method to predict species distribution facing a changing world, our results may be considered as a call for attention for researchers. Accurate and reliable predictions are increasingly needed and, as it seems to be, models which combine climate and biotic variables are closer to reality (Palacio and Girini 2018). This kind of data may be more easily found and applied for species with specialized habits such as those which feed mainly in one kind of food resource (Giannini et al. 2013). Furthermore, the food web control may also be a driver in model accuracy. In a bottom-up control for instance, higher levels need the lower ones but not necessarily the other way around (Vázquez et al. 2007; Giannini et al. 2013).

Sciurus aestuans bases its diet in, and is the main disperser of, *S. romanzoffiana*, which indicates high specialization between these two species (Alvarenga and Talamoni 2006). This fact may be responsible for the high contribution of *S. romanzoffiana* distribution in the prediction model of *S. aestuans* distribution. This kind of result is more easily found in strong and specialized interactions (Giannini et al. 2013). However, understanding distribution of generalist species is also needed, making the knowledge of species ecology a bottleneck in SDMs (Araújo et al. 2014). Thus, we encourage new studies that predict species distribution matching climate and biotic predictors even though it is difficult to access this kind of data for some species.

There are different ways to establish conservation planning (Margules and Pressey 2000). However, the concern about a future with climate change is a congruence (Menon et al. 2002; Duarte et al. 2013). Hence, each attempt to reach a good conservation planning needs to go through the future prediction, which is the objective of most SDMs (Thuiller et al. 2003; Synes and Osborne 2011). This way, new approaches and methodologies to accurately predict the future are appreciated. Our study may contribute to predict species distribution accurately and close to reality and, therefore, to conservation planning in a future of climatic changes.

Conclusion

We suggest an integrative approach for SDMs which include biotic and climatic predictors, as indicated by other authors (Palacio and Girini 2018; Simoes and Peterson 2018). Our results also highlight that SDMs applied without a biological purpose do not make sense. We call the researchers' attention to the application and assessment of SDMs, where each objective needs to be taken into account and the ecological sense can never be forgotten. We cannot just turn biological and ecological knowledge into numbers. Considering all this, research may be performed with different algorithms and scores but the first premise needs to be the previous knowledge of species and ecology.

References

- Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* 43(6): 1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>
- Alvarenga CA, Talamoni SA (2006) Foraging behaviour of the Brazilian squirrel *Sciurus aestuans* (Rodentia, Sciuridae). *Acta Theriologica* 51(1): 69–74. <https://doi.org/10.1007/BF03192657>
- Amori G, Chiozza F, Patterson BD, Rondinini C, Schipper J, Luiselli L (2013) Species richness and distribution of Neotropical rodents, with conservation implications. *Mammalia* 77(1): 1–19. <https://doi.org/10.1515/mammalia-2012-0050>
- Anderson RP (2017) When and how should biotic interactions be considered in models of species niches and distributions? *Journal of Biogeography* 44(1): 8–17. <https://doi.org/10.1111/jbi.12825>
- Araújo MB, Luoto M (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography* 16(6): 743–753. <https://doi.org/10.1111/j.1466-8238.2007.00359.x>
- Araújo MB, Rozenfeld A (2013) The geographic scaling of biotic interactions. *Ecography*: 1–10. <https://doi.org/10.1111/j.1600-0587.2013.00643.x>
- Araújo CB, Marcondes-Machado LO, Costa GC, Silman M (2014) The importance of biotic interactions in species distribution models: A test of the Eltonian noise hypothesis using parrots. *Journal of Biogeography* 41(3): 513–523. <https://doi.org/10.1111/jbi.12234>
- Bakkenes M, Alkemade J, Ihle F, Leemans R, Latour J (2002) Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. *Global Change Biology* 8(4): 390–407. <https://doi.org/10.1046/j.1354-1013.2001.00467.x>
- Barry S, Elith J (2006) Error and uncertainty in habitat models. *Journal of Applied Ecology* 43(3): 413–423. <https://doi.org/10.1111/j.1365-2664.2006.01136.x>
- Baselga A, Araújo MB (2009) Individualistic vs community modelling of species distributions under climate change. *Ecography* 32(1): 55–65. <https://doi.org/10.1111/j.1600-0587.2009.05856.x>
- Bordignon M, Monteiro-Filho EL (2000) O serelepe *Sciurus ingrami* (Sciuridae: Rodentia) como dispersor do pinheiro do Paraná *Araucaria angustifolia* (Araucariaceae: Pinophyta). *Arquivos de Ciências Veterinárias e Zoologia da UNIPAR*: 3.
- Diniz-Filho JAF, Mauricio Bini L, Fernando Rangel T, Loyola RD, Hof C, Nogués-Bravo D, Araújo MB (2009) Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. *Ecography* 32: 897–906. <https://doi.org/10.1111/j.1600-0587.2009.06196.x>
- Duarte CM, Losada IJ, Hendriks, IE, Mazarrasa I, Marbà N (2013) The role of coastal plant communities for climate change mitigation and adaptation. *Nature Climate Change* 3(11): 961–968. <https://doi.org/10.1038/nclimate1970>
- Elith J, Graham CH (2009) Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography* 32(1): 66–77. <https://doi.org/10.1111/j.1600-0587.2008.05505.x>

- Engler J, Rödder D, Elle O, Hochkirch A, Secondi J (2013) Species distribution models contribute to determine the effect of climate and interspecific interactions in moving hybrid zones. *Journal of Evolutionary Biology* 26(11): 2487–2496. <https://doi.org/10.1111/jeb.12244>
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24(1): 38–49. <https://doi.org/10.1017/S0376892997000088>
- Giannini TC, Chapman DS, Saraiva AM, Alves-dos-Santos I, Biesmeijer JC (2013) Improving species distribution models using biotic interactions: A case study of parasites, pollinators and plants. *Ecography* 36(6): 649–656. <https://doi.org/10.1111/j.1600-0587.2012.07191.x>
- Grinnell J (1917) The niche-relationships of the California Thrasher. *The Auk* 34(4): 427–433. <https://doi.org/10.2307/4072271>
- Guillera-Arroita G, Lahoz-Monfort JJ, Elith J, Gordon A, Kujala H, Lentini PE, McCarthy MA, Tingley R, Wintle BA (2015) Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography* 24(3): 276–292. <https://doi.org/10.1111/geb.12268>
- Guisan A, Thuiller W (2005) Predicting species distribution: Offering more than simple habitat models. *Ecology Letters* 8(9): 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>
- Haila Y, Kouki J (1994) The phenomenon of biodiversity in conservation biology. *Journal Name and volume*: 5–18. <https://www.jstor.org/stable/23735495>
- Heikkinen RK, Luoto M, Kuussaari M, Toivonen T (2007) Modelling the spatial distribution of a threatened butterfly: Impacts of scale and statistical technique. *Landscape and Urban Planning* 79(3–4): 347–357. <https://doi.org/10.1016/j.landurbplan.2006.04.002>
- Hijmans RJ (2011) Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology* 93(3): 679–688. <https://doi.org/10.1890/11-0826.1>
- Huntley B, Berry PM, Cramer W, McDonald AP (1995) Special paper: Modelling present and potential future ranges of some European higher plants using climate response surfaces. *Journal of Biogeography* 22(6): 967–1001. <https://doi.org/10.2307/2845830>
- Hutchinson G (1957) A treatise on limnology vol. II. John Wiley and Sons, New York, 1015 pp.
- Kadmon R, Farber O, Danin A (2003) A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications* 13(3): 853–867. [https://doi.org/10.1890/1051-0761\(2003\)013\[0853:ASAOFA\]2.0.CO;2](https://doi.org/10.1890/1051-0761(2003)013[0853:ASAOFA]2.0.CO;2)
- Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M (2017) Climatologies at high resolution for the earth's land surface areas. *Scientific Data* 4(1): e170122. <https://doi.org/10.1038/sdata.2017.122>
- Liu C, White M, Newell G (2011) Measuring and comparing the accuracy of species distribution models with presence-absence data. *Ecography* 34(2): 232–243. <https://doi.org/10.1111/j.1600-0587.2010.06354.x>
- Lobo JM, Lumaret JP, Jay-Robert P (2002) Modelling the species richness distribution of French dung beetles (Coleoptera, Scarabaeidae) and delimiting the predictive capacity

- of different groups of explanatory variables. *Global Ecology and Biogeography* 11(4): 265–277. <https://doi.org/10.1046/j.1466-822X.2002.00291.x>
- Loiselle BA, Howell CA, Graham CH, Goerck JM, Brooks T, Smith KG, Williams PH (2003) Avoiding pitfalls of using species distribution models in conservation planning. *Conservation Biology* 17(6): 1591–1600. <https://doi.org/10.1111/j.1523-1739.2003.00233.x>
- Margules CR, Pressey RL (2000) Systematic conservation planning. *Nature* 405(6783): 243–253. <https://doi.org/10.1038/35012251>
- Margules C, Usher M (1981) Criteria used in assessing wildlife conservation potential: A review. *Biological Conservation* 21(2): 79–109. [https://doi.org/10.1016/0006-3207\(81\)90073-2](https://doi.org/10.1016/0006-3207(81)90073-2)
- Meier R, Dikow T (2004) Significance of specimen databases from taxonomic revisions for estimating and mapping the global species diversity of invertebrates and repatriating reliable specimen data. *Conservation Biology* 18(2): 478–488. <https://doi.org/10.1111/j.1523-1739.2004.00233.x>
- Meier ES, Edwards Jr TC, Kienast F, Dobbertin M, Zimmermann NE (2011) Co-occurrence patterns of trees along macro-climatic gradients and their potential influence on the present and future distribution of *Fagus sylvatica* L. *Journal of Biogeography* 38(2): 371–382. <https://doi.org/10.1111/j.1365-2699.2010.02405.x>
- Meineri E, Skarpaas O, Vandvik V (2012) Modeling alpine plant distributions at the landscape scale: Do biotic interactions matter? *Ecological Modelling* 231: 1–10. <https://doi.org/10.1016/j.ecolmodel.2012.01.021>
- Menon S, Hansen J, Nazarenko L, Luo Y (2002) Climate effects of black carbon aerosols in China and India. *Science* 297(5590): 2250–2253. <https://doi.org/10.1126/science.1075159>
- Palacio FX, Girini JM (2018) Biotic interactions in species distribution models enhance model performance and shed light on natural history of rare birds: A case study using the straight-billed reedhaunter *Limnocittes rectirostris*. *Journal of Avian Biology* 49(11): e01743. <https://doi.org/10.1111/jav.01743>
- Paschoal M, Galetti M (1995) Seasonal food use by the neotropical squirrel *Sciurus ingrami* in southeastern Brazil. *Biotropica* 27(2): 268–273. <https://doi.org/10.2307/2389006>
- Pearce J, Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling* 133(3): 225–245. [https://doi.org/10.1016/S0304-3800\(00\)00322-7](https://doi.org/10.1016/S0304-3800(00)00322-7)
- Pearson RG, Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? *Global Ecology and Biogeography* 12(5): 361–371. <https://doi.org/10.1046/j.1466-822X.2003.00042.x>
- Pearson RG, Raxworthy CJ, Nakamura M, Townsend Peterson A (2007) Predicting species distributions from small numbers of occurrence records: A test case using cryptic geckos in Madagascar. *Journal of Biogeography* 34(1): 102–117. <https://doi.org/10.1111/j.1365-2699.2006.01594.x>
- Peterson AT, Papes M, Soberón J (2008) Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling* 213(1): 63–72. <https://doi.org/10.1016/j.ecolmodel.2007.11.008>

- Schweiger O, Heikkinen RK, Harpke A, Hickler T, Klotz S, Kudrna O, Kühn I, Pöyry J, Settele J (2012) Increasing range mismatching of interacting species under global change is related to their ecological characteristics. *Global Ecology and Biogeography* 21(1): 88–99. <https://doi.org/10.1111/j.1466-8238.2010.00607.x>
- GBIF (2020) *Sciurus aestuans* Linnaeus, 1766. GBIF Backbone Taxonomy. <https://doi.org/10.15468/39omei> [accessed on 2021-04-16]
- Segurado P, Araujo MB (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography* 31(10): 1555–1568. <https://doi.org/10.1111/j.1365-2699.2004.01076.x>
- Simoes MVP, Peterson AT (2018) Importance of biotic predictors in estimation of potential invasive areas: The example of the tortoise beetle *Eurypedus nigrosignatus*, in Hispaniola. *PeerJ* 6: e6052. <https://doi.org/10.7717/peerj.6052>
- Soares KP (2020) *Syagrus*. Flora do Brasil 2020. Jardim Botânico do Rio de Janeiro. <http://floradobrasil.jbrj.gov.br/reflora/floradobrasil/FB15745> [accessed on 23/11/2019]
- Soberón J, Nakamura M (2009) Niches and distributional areas: Concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences of the United States of America* 106(Supplement 2): 19644–19650. <https://doi.org/10.1073/pnas.0901637106>
- Synes NW, Osborne PE (2011) Choice of predictor variables as a source of uncertainty in continental-scale species distribution modelling under climate change. *Global Ecology and Biogeography* 20(6): 904–914. <https://doi.org/10.1111/j.1466-8238.2010.00635.x>
- Thuiller W, Araújo MB, Lavorel S (2003) Generalized models vs. classification tree analysis: Predicting spatial distributions of plant species at different scales. *Journal of Vegetation Science* 14(5): 669–680. <https://doi.org/10.1111/j.1654-1103.2003.tb02199.x>
- Vázquez DP, Melián CJ, Williams NM, Blüthgen N, Krasnov BR, Poulin R (2007) Species abundance and asymmetric interaction strength in ecological networks. *Oikos* 116(7): 1120–1127. <https://doi.org/10.1111/j.0030-1299.2007.15828.x>
- Wisz MS, Pottier J, Kissling WD, Pellissier L, Lenoir J, Damgaard CF, Dormann CF, Forchhammer MC, Grytnes JA, Guisan A, Heikkinen RK, Høye TT, Kuhn I, Luoto M, Maiorano L, Nilsson MC, Normand S, Ockinger E, Schmidt NM, Termansen M, Timmermann A, Wardle DA, Aastrup P, Svenning JC (2013) The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological reviews of the Cambridge Philosophical Society* 88: 15–30. <https://doi.org/10.1111/j.1469-185X.2012.00235.x>
- Zimmermann NE, Edwards Jr TC, Graham CH, Pearman PB, Svenning JC (2010) New trends in species distribution modelling. *Ecography* 33(6): 985–989. <https://doi.org/10.1111/j.1600-0587.2010.06953.x>

Supplementary material 1

Occurrence data – *Syagrus romanzoffiana*

Authors: Gabriel Preuss, André Andrian Padial

Data type: occurrence

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/neotropical.16.e64892.suppl1>

Supplementary material 2

Occurrence data – *Sciurus aestuans*

Authors: Gabriel Preuss, André Andrian Padial

Data type: occurrence

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/neotropical.16.e64892.suppl2>

Supplementary material 3

Code used to perform all analysis

Authors: Gabriel Preuss, André Andrian Padial

Data type: code

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/neotropical.16.e64892.suppl3>

Supplementary material 4

Tables S1, S2, Figure S1

Authors: Gabriel Preuss, André Andrian Padial

Data type: measurements

Explanation note: **Table S1.** Values of variables importance for each simulation inside the ensemble. **Table S2.** Values of performance scores and presence pixels for each simulation inside ensemble. **Figure S1.** Variation in presence pixels among the models of the cross validation (n=10), where: x axis represents the model type and y axis represents the number of presence pixels in the map.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/neotropical.16.e64892.suppl4>